

Laypeople's Views on Decision Making in the Health Professions

A Senior Honors Thesis

Presented in partial fulfillment of the requirement for graduation with distinction in Psychology
in the undergraduate colleges of the Ohio State University

By

Stephanie de Oliveira

The Ohio State University

June 2009

Project Advisor: Dr. Hal Arkes, Department of Psychology

Abstract

Although actuarial judgment (AJ) is often superior to clinical judgment (CJ), it is underused in the health fields for decision making. One potential reason is that people would not like to be treated as statistics. This study investigates laypeople's views about the use of AJ versus CJ. Manipulations include domain – a medical context or a psychological context – and perceived difficulty of a diagnosis on laypersons' preference for either method. It is hypothesized that laypeople are biased against AJ in general. However there are many potential effects of varying context and difficulty. Therefore several possibilities are discussed.

82 Research Experience Program students participating for credit were randomly assigned to two groups in which all materials were presented in a medical or psychological context. Participants read unbiased definitions of AJ and CJ, then read scenarios where easy or difficult diagnoses were made by physicians or psychologists. After each scenario, participants indicated how appropriate it was for the professional to use CJ and AJ. Final questions further explored participants' comfort levels with each method, and their views of health professionals and actuarial tool development.

Results indicated significant interactions between style and difficulty on the rated appropriateness of AJ and CJ. After reading the easy scenario, participants were more in favor of CJ, but after the difficult scenario they rated AJ as more appropriate. Significant interactions between style and domain revealed that CJ was perceived as more appropriate in a psychological context, but appropriateness inverted in the medical context. In general, participants were more comfortable with CJ in both the medical and psychological groups. Finally, people perceived physicians as more knowledgeable than psychologists, and their decision making to be more rational. Possible factors influencing these views are discussed in the paper.

Laypeople's Views on Decision Making in the Health Professions

Two prominent means of integrating information to yield judgments, predictions and decisions are the actuarial method (also known as mechanical or statistical judgment) and the clinical method. The former method involves treating data in a purely mathematical and empirical manner, as exemplified by the use of regression equations and computerized decision aids. The latter technique relies on human judgment which may be comprised of intuition, expertise, and any mental exercise that yields a decision.

To understand these two methods as discussed in this paper, it is necessary to distinguish between the information-collecting process and the integration of information. The nature of the collection process does not define whether the overall judgment is actuarial or clinical; it may require the use of diverse activities in order to gather the required information for the decision. Holt (1970) outlines six steps that comprise a typical decision making journey for a human judge. The judge must decide on a criterion – what it is one wants to predict – and he must also choose predictors. Measuring instruments are then selected, and the predictors are given empirical trials to test their quality. Data are then gathered and scored if necessary. Finally, scores must be combined somehow to make the final judgment or prediction. Holt emphasizes that even a method dubbed “actuarial” may have strong clinical influence up to the final step. For the purposes of this document, whether a method is termed clinical or actuarial will heavily depend on what method is used to *integrate* the collected data. Therefore, an actuarial method does not preclude the use of clinical judgment in an earlier step – it is conceded that humans are important even in a mathematical method. Likewise, a clinical method does not preclude the use of an actuarial tool in an earlier step.

Much controversy exists over which is the optimal method to be used in the health fields. Doctors, nurses, and psychologists are constantly faced with important judgments to make such as diagnoses, choosing treatments, and predicting the course of illnesses in different patients. An immense amount of information must be examined to make such judgments, and the quality of their judgments is crucial to the well-being and survival of many human beings. Controversy over the optimal judgment method does not only involve comparing the technical quality of the methods (accuracy and reliability, for example), but it also requires comparison of relative financial costs, the availability of tools for actuarial judgment, and the preferences and opinions of patients being examined.

The relative technical quality of actuarial and clinical methods has been heavily investigated. It is currently reasonable to suggest that in many situations, actuarial judgment is superior to clinical judgment (Dawes, 1979; Grove & Meehl, 1996; Grove, Zald, Lebow, Snitz, & Nelson, 2000). This is likely true in part due to the biases and inappropriate heuristics that the unaided instinctual mind falls for. For example, one's retrieval of information from memory can be biased. Reyes and others found that by only manipulating the vividness of either the prosecution evidence or the defense evidence in a court scenario they could affect the jurors' verdicts (Reyes, Thompson, & Bower, 1980). This effect was stronger 48 hours later, when the jurors relied more on memory to make their decision. Another common mistake is that people judge randomness by how similar a sequence is to their stereotype of randomness, which usually has too much variation for the small samples examined. Therefore, people often suppose there is a cause in events that are truly random but do not conform to their stereotype of randomness. For example, many think that the "hot hand" is behind a series of successful shots in basketball, when a random streak of success is not all that unlikely to happen (Gilovich, Vallone, &

Tversky, 1985). As a final example, Chapman and Chapman (1967) found that when non-clinicians observed 45 patient drawings of people randomly paired with diagnoses, participants detected non-existent relationships between drawing features and diagnoses. This suggests that people often see covariation when there is none, and that by computing the actual correlation between, say, the drawing's eye-size and paranoid diagnoses, this error could be avoided. Human intuitive judgment is contaminated with a host of flaws, and this ought to provide motivation for seeking superior methods of judgment in critical situations such as the emergency room.

An actuarial tool to aid in a prediction or judgment task can be developed by using objective statistics such as the relationship particular symptoms have with a disorder, the base rates of the disorder, and the import of medical history on the diagnosis. The same information can be collected every time, the same weights placed on each piece of data, and the same integration method can be used. This consistency paired with careful development of the tool itself can render better predictive accuracy than an unaided mind.

Yet such tools are underused (Böckenholt & Weber, 1992). Suggested reasons for under-use include low awareness of the existence of such methods, low motivation to change from the status-quo, and perceived inadequacies of the actuarial tools (Böckenholt & Weber, 1992; Corey & Merenstein, 1987). In the medical setting specifically, the use of certain computerized tools may offend physicians' sense of professional autonomy and strip medicine of its "artistic" tone (Kaplan, 2000).

Problems with assuming the superiority of actuarial methods in clinical, psychological settings are worth addressing. It appears reasonable to assert that since *general* human judgment is inferior to actuarial judgment, then judgments specifically made in a clinical setting must also be inferior. How could one protest that this domain of judgment has a special barrier against

general flaws such as biases? Yet the assertion that actuarial judgment is superior to clinical judgment in the psychologist's office remains very difficult to empirically test at this point in time. In medical settings, "gold standards" are available to test the human predictions against hard, conclusive evidence – a physician may diagnose a broken ankle and then use an x-ray to confirm the diagnosis. But it is unrealistic to say that a clinician may diagnose bipolar disorder in a patient and then use a structured interview to confirm the diagnosis. Given the leap taken between clinical assumptions and "reality", and science's present inability to confirm the "reality" behind many mental illnesses, it remains difficult to empirically compare the performance of a clinician against the performance of an actuarial test in a clinical context.

This obstacle by no means allows professionals to settle for the status-quo. Other facts still speak in favor of the actuarial method. First, actuarial methods yield perfectly re-producible and consistent judgments. Second, it been demonstrated that even improperly constructed linear models perform better than humans in many judgment tasks (Dawes, 1979). As for the low availability of usable actuarial tools, they can certainly be developed, as a few have been already (Corey & Merenstein, 1987; Medow et al., 2001). It is even possible to make mathematical models of human judges that outperform the judges in prediction tasks (Wiggins & Kohen, 1971). Finally, it is also likely that the financial cost of developing and using an actuarial decision aid is lower in the long run than paying the professional for the hours spent at work integrating information (Grove & Meehl, 1996; Grove et al., 2000). Taking into account the *general* superiority of actuarial methods, the possibility of developing more tools, and the lower cost of doing so, it might be concluded that using actuarial methods is worthwhile even in a clinical situation.

Still another objection has been raised and presented as a reason for the low adoption of decision aids/ actuarial tools. It is maintained that treating unique individuals as statistics is “dehumanizing” (Dawes, 1989; Grove & Meehl, 1996). It has also been demonstrated that patients may think less highly of health professionals who rely on decision aids, although this finding is somewhat debatable (Arkes, Shaffer, & Medow, 2007; Pezzo & Pezzo, 2006). However Dawes (1979), among others, has pointed out that although treating people as numbers may be uncomfortable, if this is indeed the best method for offering quality service then it is certainly more ethical to use it than not (Grove & Meehl, 1996).

From the perspective of a professional well-informed in this area of judgment and decision making, there should remain no question that mathematical aids ought to be used when they are available and shown to be superior to unaided human judgment. Yet is it ethical to use them if those who are paying for this judgment service are against them? Objectors to actuarial practice claim that the use is dehumanizing, and it could be considered so if clients felt offended or uncomfortable when actuarial methods were used. The research to date has not yielded copious information on what the general population's attitudes are towards actuarial practices. Previously-mentioned studies have investigated laypersons' attitudes towards professionals who use or do not use decision-aids. Arkes and colleagues (2007) conducted a series of experiments in which subjects gave physicians lower ratings for diagnostic ability if the physician used a decision aid to make a diagnosis. Pezzo and Pezzo (2006) found that when doctors' diagnoses were correct, those who used aids were viewed less positively by subjects. Yet when diagnoses were incorrect, those who used aids were viewed less negatively than those who did use aid. Laypeople's views about physicians have been studied, but their views on aid-use itself and views about psychologists who use aids are not clear.

It is possible that a client would think less of a health professional for using a decision aid and still support its use. It may be that it is too difficult to avoid the discounting principle which proposes that when multiple sources could be causing an event, people may assign responsibility primarily to one and discount the other (Kelley, 1973). Therefore the client would attribute a correct diagnosis primarily to the aid, consequently making the rating of the professional lower due to discounting. Laypersons' views of actuarial method use should be further investigated, particularly in situations where this method performs favorably, or when no information about performance is given.

The current study is designed to discover if, in general, laypersons have an aversion to actuarial decision making methods as they are used in clinical or medical settings, particularly for diagnostic purposes. It is hypothesized that people will generally be biased against actuarial decision making. Although aid use might not be negatively rated, it is hypothesized that clinical methods are preferred. It is also suggested that differences between the two settings will influence laypeople's bias strengths. A medical test's validity can be compared to known outcomes, whereas a psychological test's performance is usually more difficult to assess. People may also perceive the two fields differently – a medical decision might seem more “concrete,” based on hard physical facts. A psychological analysis may appear more subjective and intuitive, given the historical psychoanalytic stereotype involving inkblot tests and couches. Due to the potential differences in views, laypersons may feel that statistical procedures are more inappropriate to use in a psychological setting than in a medical setting. In other words, laypeople may overestimate how objective medical decision making is and overestimate how subjective mental assessment is, and an objective way of integrating data may seem more appropriate in the former setting.

An opposing argument predicts opposite results. There is evidence that medical professionals – physicians in particular – may be endowed with an “aura of omniscience” (Arkes et al, 2007, p. 196). As previously cited, they are viewed as more skilled if they do not use decision aids to make diagnoses. Because of this phenomenon, decision aids may seem unnecessary altogether. A discrepancy in the “auras” of physicians and psychologists as perceived by laypeople may yield a discrepancy in ratings of the appropriateness of decision-aid use. In the context where omniscience bestowed on the professional is greater – possibly the medical context – decision aids would be deemed as less necessary and therefore participants would display a stronger bias against them. Because the roles of professional objectivity and omniscience in the general population’s perception of health professionals are currently unclear, support of either hypothesis is possible.

The present study also aims to study effects of perceived difficulty of the diagnosis on one’s preference for actuarial or clinical methods. It is hypothesized that an easy diagnosis in both settings will yield mild or moderate preferences for clinical decision making, and this effect is expected to interact with context although direction is unclear. A difficult diagnosis may strengthen the preference for a clinical diagnosis, as it is the prevalent method and more familiar to most people. Yet again, an opposing hypothesis is also plausible: the more difficult diagnosis may be interpreted as a suggestion that the professional “needs help.” In this situation, participants would be more willing to accept actuarial tools rather than less willing.

The participants in this study shall receive no statements about relative efficacies of clinical versus actuarial diagnoses, as the dependent measures test for biases. If participants have no such data from which to make a judgment about method preference, the only resource from which they may make a decision is their affective response to the definitions. If there is no

affective reason to choose one method over the other, i.e. there is no aversion toward mathematical methods, then the net results for a group should indicate no bias. Significant deviations from such a balance would indicate preference bias stemming from considerations other than empirically-tested method efficacy. This bias should be taken into consideration when discussing strategies to implement quality actuarial-tool-use in professional practice.

Method

Participants

182 participants were recruited using Ohio State's Research Experience Program (REP). They were undergraduate students enrolled in Psychology 100, for which they received credit after participating. Materials were presented on PC's using MediaLab and variables were collected via the same medium.

Design

The students were randomly placed into two groups, with group placement determining whether a participant will receive materials pertaining to medical issues only or psychological issues only. Within each group, every participant will receive two scenarios in which a health professional must make one difficult diagnosis and one easy diagnosis. A given participant's easy and difficult scenarios will be presented within the same context, and counterbalanced in order. After each scenario, participants are asked in two questions to rate how appropriate it would be if the professional used clinical or actuarial decision making. A manipulation check for difficulty was included after each scenario in which participants were asked to fill in a blank: "I believe that diagnosing this illness is ____." Five options included "very easy" on one end (coded as 1) and very difficult on the other end (coded as 5). After both scenarios were finished, additional questions were presented to measure participant attitudes about health professionals.

Thus, manipulating professional domain, diagnostic difficulty, scenario order and decision making style, yields a 2 (domain, between subjects) x 2 (difficulty, within subjects) x 2 (order, between subjects) x 2 (style, within subjects) design.

Materials

The computer materials first provided an introduction to the study topic and objective definitions of clinical and actuarial decision making. Within the introduction, a reference was made to easy and difficult decisions, i.e. health professionals “may need to examine many symptoms and tests to diagnose an illness. This could be either a very easy or very difficult task.” This was intended to help diminish contrast effects when participants read about two different scenarios later in the study. That is, participants would bear in mind the range of difficulties that a decision could have and then mentally calibrate the scale. All participants received identical definitions written on paper as a reference for later questions. The definitions were as follows:

In the clinical method, the decision-maker combines or processes information in his or her head. He or she can rely on mental tools like expertise, intuition, and experience.

Note that it's acceptable to use textbooks or journals to look up information – the main point is that the decision is made by the person using his or her own mental power.

In the actuarial method, a statistical procedure is used to reach the decision. This method often uses mathematical tools that were developed from studying empirically established relationships between events and conditions. Again, it's acceptable to use textbooks or journals to look up information. The main point is that the decision maker, who must make the final decision, relies heavily on the mathematical formula's output.

After reading the objective definitions of each method, all groups were given two scenarios in which easy or difficult diagnoses were to be made by psychologists or physicians. The diseases and disorders selected were fairly uncommon and were presented as rather debilitating, though not fatal. The former tactic served to decrease the possibility that the participants would superimpose their own knowledge about diagnostic difficulty onto the illness, and the latter tactic ensured that the decision would not be taken lightly. The diagnostic difficulty manipulation was executed by varying whether there were idiosyncratic symptoms that facilitated diagnosis, whether the illness was clearly distinguishable from similar conditions, and whether a final statement told the participant directly that the diagnosis was easy or difficult. Illnesses chosen for the medical groups were lupus and Crohn's disease; those for the psychological group were schizoaffective disorder and borderline personality disorder. The presentation was designed so that each disorder could be displayed as either the easy or the difficult ailment – the name itself had no bearing on whether the illness was presumed to be easy or difficult to diagnose. The following example is the difficult-diagnosis medical scenario:

Imagine that a close friend is seeing a physician. Your friend has been complaining of fatigue and has been feeling generally ill for a few weeks. The physician suspects that your friend may have Lupus. You receive the following information about the disease:

Lupus resembles a variety of illnesses, and it is difficult to tell them apart.

The symptoms are vague.

A misdiagnosis could significantly reduce the quality of life for your close friend. Because of this information, you and your friend feel that it is extremely important that the correct diagnosis is made. Getting the correct diagnosis is very difficult, so you and your friend want to choose the best decision making method possible.

The physician has collected information about symptoms and family history, and can either give your friend a diagnosis based on expertise (clinical) or on output from a computer program (actuarial). You wonder, “What’s a good way to integrate the information to get a diagnosis?”

Next, two questions plus non-numerical answer scales were presented: “In this situation, basing a diagnosis on the physician’s clinical judgment is ____” and “In this situation, basing the diagnosis on the actuarial computer output is ____”. The scales’ seven points were: Very inappropriate, moderately inappropriate, slightly inappropriate, neutral, slightly appropriate, moderately appropriate, and very appropriate. The responses were coded with 1 being “very inappropriate” and 7 being “very appropriate”. After giving responses regarding decision making modes, participants were asked how difficult they believed the diagnosis to be. This manipulation check offered five response options, ranging from very difficult (coded as 5) to very easy (coded as 1).

After completing both scenarios and responses, all participants were asked two more fill-in-the-blank, multiple-choice questions. The sentences to fill in for the medical group were, “I would feel ____ seeing a physician who made diagnoses primarily based on clinical judgment” and “I would feel ____ seeing a physician who made diagnoses primarily based on actuarial tools’ output.” For the psychology group, the word “psychologist” replaced “physician.” Four answer choices included: Very comfortable (coded as 1), comfortable, uncomfortable, and very uncomfortable (coded as 4).

Subsequently, participants were asked to provide their levels of agreement or disagreement to three statements: “It is a good idea to develop good actuarial tools (i.e. ones that are more accurate than professionals alone), even if it takes time and money,” “Actuarial

methods are not useful because they yield results based on information about large groups of people, whereas professionals make judgments about unique individuals,” and “Using actuarial diagnostic methods would make the relationship between professionals and clients less warm and personal.” Seven response options range from “strongly disagree” (coded as 1) to “strongly agree” (coded as 7). Participants then were asked to agree or disagree with four sentences, using the same seven-point response scale as was just mentioned. The last word in each of the four sentences was varied: “In general, the way that most physicians/psychologists make decisions is intuitive/ rational/ objective/ subjective.”

Finally, information was gathered to elucidate questions generated by the context hypotheses. Participants described their health professional's decision making methods by rating the methods on how intuitive, rational, subjective, and objective they are. To measure perceived omniscience, participants rated their respective professionals on a five-point word scale ranging from very knowledgeable to very unknowledgeable. A final question asked participants whether they had had previous exposure to the actuarial/clinical issue, and to describe such exposure in an open-ended reply.

Results

Eighty-two participants completed the study for class credit. Analyses were conducted twice, once including all participants and once excluding twelve who rated the difficult scenario as easier than the easy scenario, leaving 70 participants. Unless noted, results noted as significant were significant in both analyses, and data presented in this section are drawn from the 82-student analyses. No exclusions due to prior knowledge were made, because no students reported a strong background in judgment and decision making, and seven participants who expressed some familiarity with the illnesses in the study passed the manipulation check.

Before more extensive analysis of the data, a manipulation check was conducted to ensure that the difficulty level of each scenario was accurately rated by participants. We used a repeated measures analysis of variance (ANOVA) from the Statistical Package for the Social Sciences (SPSS), version 17.0, on the difficulty ratings for the easy scenario and the difficulty ratings for the difficult scenario. Difficulty was a within-subject factor (1= difficult, 2= easy), as each participant read a difficult and easy scenario. Between-subject factors were domain (1= psychological, 2= medical) and order (1= easy first, 2= difficult first). Scenario difficulty alone had a significant effect on difficulty ratings, $F(1, 78) = 27.58, p < .01$. The more difficult scenario ($M = 3.8, SE = .10$) was rated as significantly more difficult than the easy scenario ($M = 2.99, SE = .12$). Note that the scale ranged from 1 (very easy) to 5 (very difficult).

A repeated measures test was performed to examine differences between appropriateness ratings of a professional's decision making method. Thus, the dependent variable was appropriateness, as influenced by two within-subjects factors: style (within subjects: actuarial and clinical) and difficulty of the decision (within subjects: easy and difficult), plus two between-subjects factors: domain (between subjects: psychological and medical) and order of scenarios (between subjects: easy first and difficult first).

Difficulty was a significant influence on the appropriateness ratings $F(1,78) = 8.42, p < .01$, with higher appropriateness ratings given to easy as opposed to difficult cases ($2.67, SE = .11$ vs. $2.96, SE = .10$). However this result was qualified by an interesting two-way interaction between style and difficulty, $F(1, 78) = 11.50, p < .01$. When the participants were reading about an easy scenario, clinical decision making was viewed as more appropriate than actuarial decision making ($M = 2.19, SE .16$ vs. $M = 3.16, SE = .20$) – recall that lower means signify more appropriateness. When participants read about a difficult decision scenario, actuarial

decision making was rated as more appropriate ($M = 2.79$, $SE .18$, vs. $M = 3.13$, $SE = .20$). Post-hoc tests including all subjects showed a significant difference in ratings only in the easy scenario in which subjects preferred clinical judgment over actuarial; $F(1, 80) = 5.50$, $p < .05$. When the data from the 12 subjects who failed the manipulation test were removed, the difference in ratings in the easy scenario in which subjects preferred clinical judgment over actuarial remained significant; $F(1, 68) = 10.14$, $p < .001$. (All post-hoc tests were collapsed over order so as to simplify the calculations.)

Insert Figure 1 here

There was a two-way interaction between the professional's domain and the style, $F(1, 78) = 6.63$, $p < .05$. For the participants reading about a psychologist, clinical decision making was rated as more appropriate than actuarial decision making ($M = 2.32$, $SE = .22$ vs. $M = 3.23$, $SE = .22$). Again, lower numbers convey more appropriateness. For the participants reading about a physician, actuarial decision making was rated as more appropriate than clinical decision making ($M = 2.71$, $SE = .21$ vs. $M = 3.00$, $SE .21$). Post-hoc tests including all subjects revealed significant differences between appropriateness ratings in the psychological group; clinical judgment was significantly preferred over actuarial judgment, $F(1,80) = 5.99$, $p < .05$. This difference remained when the 12 participants were removed; $F(1, 68) = 6.18$, $p < .05$. ($M = 2.32$, $SE = .24$ vs. $M = 3.22$, $SE = .23$).

Insert Figure 2 here.

Another repeated measures test was used to check for a difference between participants' comfort levels. Comfort ratings, from 1 as most comfortable and 4 as most uncomfortable, could once again be affected by style, domain, and order, but difficulty was not a factor in this question, because it was measured after both the easy and difficult cases had been presented. In

the analyses for all participants, no effect of order was found. However, excluding the 12 revealed a main effect of order: $F(1, 66) = 4.13, p < .05$. The mean for those who read the easy scenario first was 2.42, $SE = .08$. Those reading the difficult scenario first had a mean response of 2.18, $SE = .09$. Order entered into no interactions, so this factor will not be discussed further.

Further analyses excluding the 12 participants revealed a main effect of decision style, $F(1,66) = 7.46, p < .01$, with participants being more comfortable with clinical rather than actuarial judgment ($M = 2.14, SE = .07$ vs. $M = 2.46, SE = .09$).

There was an interaction between style and domain, $F(1, 78) = 5.94, p < .05$. For those who were in the psychological group, participants were more comfortable with clinical judgment ($M = 1.98, SE = .10$) than actuarial ($M = 2.60, SE = .12$). Post-hoc tests revealed this difference in comfort level to be significant, $F(1, 80) = 19.10, p < .01$. Those in the medical group were also more comfortable with clinical judgment ($M = 2.27, SE = .09$) than with actuarial ($M = 2.37, SE = .12$), but this difference in comfort ratings was not significant.

Insert Figure 3 here.

A few post-scenario responses differed between domains. Domain had a significant effect on ratings of the health professionals' knowledge, $F(1, 80) = 6.20, p < .05$. Physicians were rated as more knowledgeable ($M = 1.37, SE = .08$) than psychologists ($M = 1.63, SE = .08$), although both were considered knowledgeable. Domain had an effect on how rational the health professional's decision making was considered, $F(1, 80) = 4.16, p < .05$. Physicians' decision making ($M = 2.19, SE = .15$) was considered more rational than psychologists' ($M = 2.63, SE = .15$). Finally, those in the medical condition were more strongly in favor of developing actuarial tools than those in the psychological group ($M = 2.12, SE = .14$ vs. $M = 2.61, SE = .14$), $F(1, 81) = 6.08, p < .05$. In the response scales for these questions, values closer to 1 indicated stronger

agreement and values approaching 7 indicated more disagreement. Other post-scenario questions' answers did not significantly differ according to domain, difficulty, or order. For the "usefulness" statement, the overall mean response was 3.28, $SD = 1.40$. The statement about professionals' warmth had an overall mean response of 2.80, $SD = 1.37$.

Discussion

These results are informative with regard to laypeople's perspectives on decision making methods. The hypotheses must be revisited in light of the results, and future studies for further clarification ought to be considered. Firstly, it was discovered that participants find clinical decision making more appropriate in psychological contexts, and in medical contexts actuarial decision making is rated as more appropriate. The differences between appropriateness ratings within each domain, as analyzed by post-hoc tests, revealed a significant difference in ratings in the psychological group.

Such inequality in appropriateness ranking, affected by style and domain, is not a result of any biasing information presented in the study. The definitions of clinical and actuarial decision making were not crafted to favor either as the normative method. Also, participants' responses were probably not influenced either by previous experience with or knowledge of judgment and decision making research. The final question queried them on such experience and knowledge, and no one described any experience that was substantially above average. Therefore, were there no other biasing agents present, the participants would have rated each form of decision making as equally appropriate or inappropriate. Yet results show that some sort of bias is present.

One possible source of bias is the presence of stereotypes about the psychological and medical professions. This relates to the hypothesis that medical decision making is viewed as

more objective than psychological decision making, rendering objective tools more appropriate. Some measures in the study support this explanation. Participants in the medical condition were significantly more strongly in favor of developing actuarial tools for professionals than those in the psychological condition. Doctors were also rated as more knowledgeable than psychologists, although this says little about their decision making style. However, participants were directly asked to rank how objective, subjective, rational, and intuitive their professional's decision making was. Answers did not differ significantly between groups except for responses to the "rational" question. In slight support of the hypothesis, physicians' decision making was rated as more rational than that of psychologists, but the hypothesis would have been more strongly supported had the medical group also rated physicians as more objective or psychologists as more subjective/intuitive.

The partial support for the stereotype difference hypothesis formulated here is consonant with findings summarized by Mills (2009), who reports that although most Americans view psychology favorably, they do not consider it a hard science. It is viewed as less valuable than medicine, and many underestimated the time it takes to get a license as a psychologist (thus explaining the difference in "knowledgeable" ratings). Finally, when asked how much they "associate the practice of psychology with scientific techniques and practices," 49% of respondents answered "somewhat" whereas only 22% answered "a great deal" (Mills, 2009, p. 29). While this study did not always directly compare beliefs about psychology with those about medicine, it is reasonable to conclude that some differences do exist and that they may be informing the way laypeople rank the appropriateness of actuarial and clinical methods.

The similar measure of comfort with each style in both domains can not be entirely explained by differences in stereotypes, as these measures yielded different results than the

appropriateness measures. The medical group responded with no significant difference in their comfort ratings, but the psychological group's responses showed significantly more comfort with clinical judgment. Comfort ratings may be a more affect-driven response, as one's comfort is not only influenced by cognitive judgments but also by emotions and physical state. Interestingly, both groups reveal conflicts in their responses between the appropriateness and comfort measures. Psychological participants are significantly more comfortable with clinical judgment but find it more appropriate than actuarial. Medical participants are more comfortable with clinical judgment yet rated actuarial as more appropriate, although their style ratings within each measure were not significantly different. This conflict may be indicative of a higher level of thinking on the part of the participants. On an emotional level, a situation may be uncomfortable; situations like setting a broken bone or humanely killing a suffering pet all yield great discomfort. Yet in many such situations, taking the less comfortable course of action is viewed as more appropriate. The same principle may be at work in this study.

The presented hypotheses also suggested reasons why people may prefer one form of decision making over another as difficulty of the decision increases. Results revealed higher appropriateness rankings for the actuarial method in the difficult scenario, but higher rankings for the clinical method in the easy scenario. However, the difference was only significant in the easy scenario. One potential reason for this is that participants may have interpreted the difficult decision as one for which the professional needed help in making a diagnosis. Such help in the form of an actuarial decision aid was deemed more appropriate than reliance on clinical judgment alone.

Another potential reason for the greater appropriateness of actuarial judgment in the difficult scenario is participants' beliefs about the quality of unaided human judgment. While

professionals across groups were generally rated as very knowledgeable, and physicians especially so, the person who perceives a decision to be difficult ranked the actuarial method as more appropriate. This suggests that if a professional's "aura of omniscience" is challenged, laypeople may be quite willing to accept the use of decision-aids and even support them.

In the real world, it is questionable whether professionals would be willing to let patients know how difficult a judgment is. Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz, and Woloshin (2008) have suggested that there has traditionally been a sense of paternalism on doctors' part in the doctor/patient relationship. This combined with trust from patients would only contribute to the "aura of omniscience." Braddock, Edwards, Hasenberg, Laidley, and Levinson (1999) examined 3,552 decisions made by physicians, such as medication decisions and laboratory tests, and found a relationship between decision complexity and completeness of the *informed* decision making process¹. It was found that 17.2% of basic decisions were complete – i.e. the patient was properly informed and involved – but only 0.5% of the complex decisions were complete ones. Such complex decisions would fall into the "difficult scenario" of this study, where the decision was explicitly said to be difficult and where participants were more in favor of actuarial judgment! If participants do not know about uncertainties or difficulties in a decision, they are not as likely to support actuarial judgment. Therefore, professionals wishing to employ decision aids but who are concerned that this would be met with disapproval may consider disclosing the difficulty of the decision. The good news from Braddock and colleagues' paper is that in complex decision situations doctors did a significantly better job of discussing *some* elements of the decision, namely pros and cons, alternatives, uncertainties, the patient's role and patient preference. Increasing such transparency with patients

¹ The researchers listed criteria for conducting a complete informed decision. This type of decision involves the patient, who is informed about various aspects of the decision such as pros and cons, alternatives, etc.

may help them gain a more realistic view of physicians' decision making and may improve their view of decision aids.

Closely related paths for further study include elucidating what stereotypes of medical and psychological professions (and professionals) are held by laypeople. Understanding in greater detail whether these beliefs differ and how they differ would help explain the interaction between domain and style found in this study. Understanding laypeople's beliefs about the quality of professionals' judgments would also help explain biases in decision making method preferences. Finally, a closely-related question asks whether comfort with actuarial judgment increases when familiarity with the style is increased.

This paper clarifies the accuracy of the objection that actuarial methods are unfavorable because they degrade unique individuals and that laypeople find aids undesirable. The final point to consider is that although biases in preferences were detected, at times favoring clinical and at times favoring actuarial, the overall appropriateness ratings for *both* methods were positive. Pezzo and Pezzo (2005) found that undergraduate students rated decision aids as less appropriate than did medical students, and that medical students were aware of this difference. The authors suggest that "perhaps it is not physicians' lack of confidence in decision aids but their awareness of patients' lack of confidence that translates into a reduced use of the aids" (p.53). Yet one must bear in mind that although the difference exists in ratings of appropriateness, in both their study and this one decision aids were rated as at least moderately appropriate. Furthermore, the use of decision aids does not necessarily decrease a patient's satisfaction with his/her medical experience (Arkes et al., 2007; see studies 2, 3, and 4). The professional's diagnostic ability may be perceived as inferior if s/he uses a decision-aid, but patients may nevertheless find the use of aid appropriate and they may be equally satisfied with the experience. Thus, such objections

ought not to be used against actuarial judgment if it is more accurate and reliable in a given context.

References

- Arkes, H. R., Shaffer, V. A., & Medow, M. A. (2007). Patients derogate physicians who use a computer-assisted diagnostic aid. *Medical Decision Making*, 27(2), 189-202.
- Böckenholt, U., & Weber, E. U. (1992). Use of formal methods in medical decision-making – A survey and analysis. *Medical Decision Making*, 12(4), 298-306.
- Braddock, C. H., Edwards, K. A., Hasenberg, N. M., Laidley, T. L., Levinson, W. (1999). Informed decision making in outpatient practice: Time to get back to basics. *Journal of the American Medical Association*, 282(24), 2313-2320.
- Chapman, L. J., & Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology*, 72(3), 193-204.
- Corey, G., & Merenstein, J. (1987). Applying the acute ischemic-heart-disease predictive instrument. *Journal of Family Practice*, 25(2), 127-132.
- Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, 34(7), 571-582.
- Gigerenzer, G., Gaissmaier, W., Kurz-Milcke, E., Schwartz, L. M., Woloshin, S. (2008). Helping doctors and patients make sense of health statistics. *Psychological Science in the Public Interest*, 8(2), 53-96.
- Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, 17, 295-314.
- Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychology, Public Policy, and Law*, 2(2), 293-323.

- Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, 12(1), 19-30.
- Holt, Robert R. (1970). Yet another look at clinical and statistical prediction: Or, is clinical psychology worthwhile? *American Psychologist*, 25(4), 337-349.
- Kaplan, B. (2000). Culture counts: How institutional values affect computer use. *M D Computing*, 17(1), 23-26.
- Kelley, H. (1973). Processes of causal attribution. *American Psychologist*, 28(2), 107-128.
- Medow, M. A., Wilt, T. J., Dysken, S., Hillson, S. D., Woods, S., & Borowsky, S. J. (2001). Effect of written and computerized decision support aids for the US Agency for Health Care Policy and Research depression guidelines on the evaluation of hypothetical clinical scenarios. *Medical Decision Making*, 21(5), 344-356.
- Mills, K. I. (2009). Getting beyond the couch: How does the general public view the science of psychology? *Monitor on Psychology*, 40(3), 28-29.
- Pezzo, M., & Pezzo, S. (2006). Physician evaluation after medical errors: Does having a computer decision aid help or hurt in hindsight? *Medical Decision Making*, 26(1), 48-56.
- Reyes, R., Thompson, W., & Bower, G. (1980). Judgmental biases resulting from differing availabilities of arguments. *Journal of Personality and Social Psychology*, 39(1), 2-12.
- Wiggins, N., & Kohen, E. S. (1971). Man vs. model of man revisited: The forecasting of graduate school success. *Journal of Personality and Social Psychology*, 19, 100-106.

Appendix A

Interaction Figures

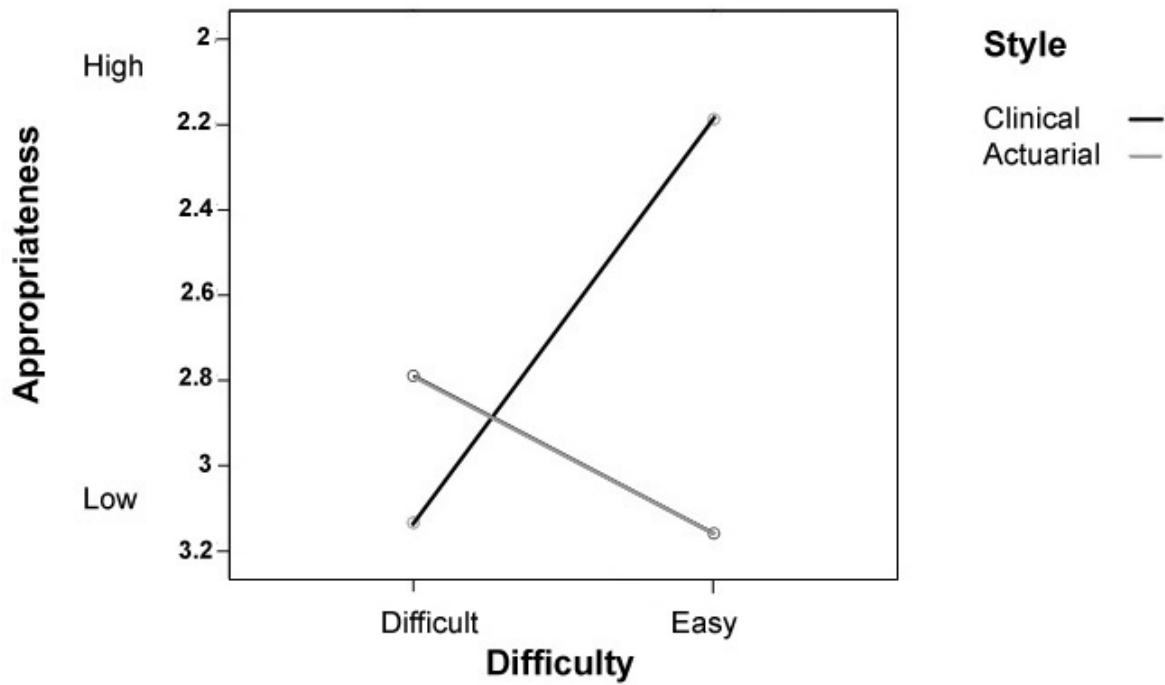


Figure 1. Interacting effects of difficulty and style on appropriateness ratings. All subjects included.

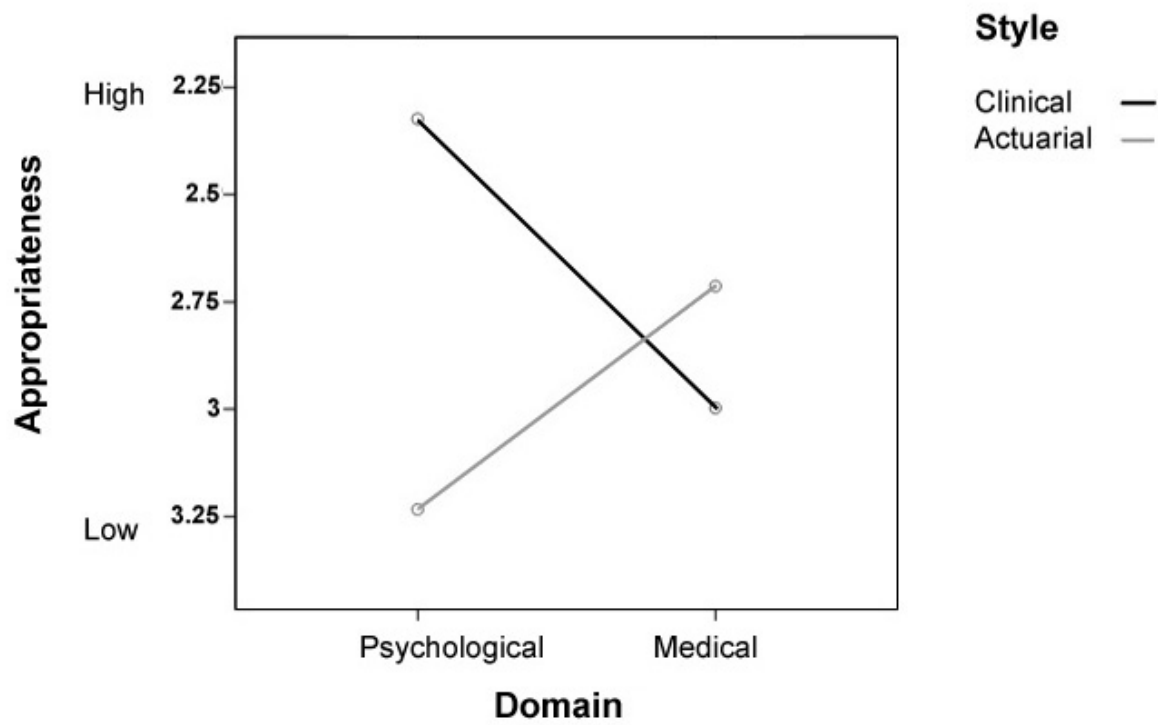


Figure 2. Interacting effects of domain and style on appropriateness ratings. All subjects included.

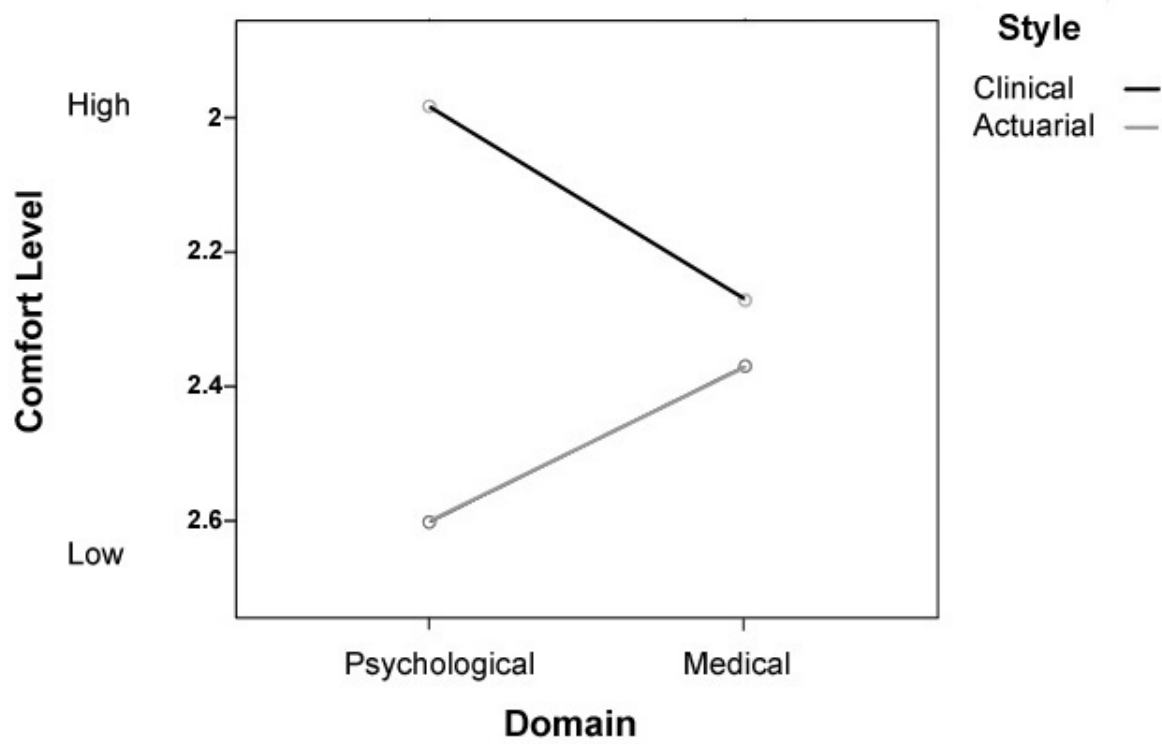


Figure 3. Interacting effects of domain and style on comfort level. All subjects included.

Appendix B

Presented Materials

Definitions

In the clinical method, the decision maker combines or processes information in his or her head. He or she can rely on mental tools like expertise, intuition, and experience. Note that it's acceptable to use textbooks or journals to look up information – the main point is that the decision is made by the person using his or her own mental power.

In the actuarial method, a statistical procedure is used to reach the decision. This method often uses mathematical tools that were developed from studying empirically established relationships between events and conditions. Again, it's acceptable to use textbooks or journals to look up information. The main point is that the decision maker, who must make the final decision, relies heavily on the mathematical formula's output.

Scenarios

Difficult scenario. Imagine that a close friend is seeing a physician/psychologist. Your friend has been (feeling generally ill) / (behaving quite strangely) for a few weeks. The physician/psychologist suspects that your friend may have Lupus/ Borderline Personality Disorder². You receive the following information about the disease/disorder:

Lupus/ Borderline Personality Disorder resembles a variety of illnesses, and it is difficult to tell them apart.

The symptoms are vague.

A misdiagnosis could significantly reduce the quality of life for your close friend.

² For the medical and psychological conditions, Crohn's disease and Schizoaffective disorder, respectively, were also at times randomly presented to participants.

Because of this information, you and your friend feel that it is extremely important that the correct diagnosis is made. Getting the correct diagnosis is very **difficult**, so you and your friend want to choose the best decision-making method possible. The physician/psychologist has collected information about symptoms and family history, and can either give your friend a diagnosis based on expertise (clinical) or on output from a computer program (actuarial). You wonder, “What’s a good way to integrate the information to get a diagnosis?”

Easy scenario. Imagine that a close friend is seeing a physician/psychologist. Your friend has been (feeling generally ill) / (behaving quite strangely) for a few weeks. The physician/psychologist suspects that your friend may have Crohn’s disease/ Schizoaffective disorder³. You receive the following information about the disease/disorder:

Crohn’s disease/ Schizoaffective disorder resembles a variety of illnesses, but it is easy to tell them apart.

The symptoms are very unique and specific.

A misdiagnosis could significantly reduce the quality of life for your close friend. Because of this information, you and your friend feel that it is extremely important that the correct diagnosis is made. Getting the correct diagnosis is very **easy**, but you and your friend still want to choose the best decision-making method possible. The physician/psychologist has collected information about symptoms and family history, and can either give your friend a diagnosis based on expertise (clinical) or on output from a computer program (actuarial). You wonder, “What’s a good way to integrate the information to get a diagnosis?”

Post-scenario questions

1) Please use the given options to complete the following sentences:

³ For the medical and psychological conditions, Lupus and Borderline Personality Disorder, respectively, were also at times randomly presented to participants.

In this situation, basing a diagnosis on the physician's clinical judgment is ____.

- ☐ Very appropriate
- ☐ Moderately appropriate
- ☐ Slightly appropriate
- ☐ Neutral
- ☐ Slightly inappropriate
- ☐ Moderately inappropriate
- ☐ Very inappropriate

2) In this situation, basing the diagnosis on the actuarial computer output is ____.

- ☐ Very appropriate
- ☐ Moderately appropriate
- ☐ Slightly appropriate
- ☐ Neutral
- ☐ Slightly inappropriate
- ☐ Moderately inappropriate
- ☐ Very inappropriate

3) I believe that diagnosing this illness is ____.

- ☐ Very difficult
- ☐ Difficult
- ☐ Neither easy nor difficult
- ☐ Easy
- ☐ Very easy

Final Questions

1) I would feel ____ seeing a physician/psychologist who made diagnoses primarily based on *clinical judgment*.

- ☐ Very Comfortable
- ☐ Comfortable
- ☐ Uncomfortable
- ☐ Very uncomfortable

2) I would feel ____ seeing a physician/psychologist who made diagnoses primarily based on *actuarial tools' output*.

- ☐ Very Comfortable
- ☐ Comfortable
- ☐ Uncomfortable
- ☐ Very uncomfortable

3) It is a good idea to develop good actuarial tools (i.e. ones that are more accurate than professionals alone), even if it takes time and money.

- ☐ Strongly agree
- ☐ Agree
- ☐ Slightly agree
- ☐ Neutral
- ☐ Slightly disagree
- ☐ Disagree
- ☐ Strongly disagree

4) Actuarial methods are not useful because they yield results based on information about large groups of people, whereas professionals make judgments about unique individuals.

- ☐ Strongly agree
- ☐ Agree
- ☐ Slightly agree
- ☐ Neutral
- ☐ Slightly disagree
- ☐ Disagree
- ☐ Strongly disagree

5) Using actuarial diagnostic methods would make the relationship between professionals and clients less warm and personal.

- ☐ Strongly agree
- ☐ Agree
- ☐ Slightly agree
- ☐ Neutral
- ☐ Slightly disagree
- ☐ Disagree
- ☐ Strongly disagree

6) In general, the way that most physicians/psychologists make decisions is intuitive:

- ☐ Strongly agree
- ☐ Agree
- ☐ Slightly agree
- ☐ Neutral
- ☐ Slightly disagree
- ☐ Disagree

- ☐ Strongly disagree

7) In general, the way that most physicians/psychologists make decisions is objective:

- ☐ Strongly agree
- ☐ Agree
- ☐ Slightly agree
- ☐ Neutral
- ☐ Slightly disagree
- ☐ Disagree
- ☐ Strongly disagree

8) In general, the way that most physicians/psychologists make decisions is subjective:

- ☐ Strongly agree
- ☐ Agree
- ☐ Slightly agree
- ☐ Neutral
- ☐ Slightly disagree
- ☐ Disagree
- ☐ Strongly disagree

9) In general, the way that most physicians/psychologists make decisions is rational:

- ☐ Strongly agree
- ☐ Agree
- ☐ Slightly agree
- ☐ Neutral
- ☐ Slightly disagree

- ☐ Disagree
- ☐ Strongly disagree

10) In general, most physicians/psychologists are:

- ☐ Very knowledgeable
- ☐ Knowledgeable
- ☐ Neither knowledgeable nor unknowledgeable
- ☐ Unknowledgeable
- ☐ Very unknowledgeable

11) Have you had any coursework on judgment and decision making, AND/OR did you already have knowledge about issues involved in this study (illnesses, decision-making)?

a) Yes

b) No

If yes, please describe any prior knowledge you already had concerning issues involved in this study:

[open-ended answer box]

Debriefing

Thank you for participating in our study. We are interested in investigating laypeople's opinions about decision-making methods in the health professions. If you would like to receive more detailed information and study results, just put a checkmark by your name on the attendance sheet as you leave. We'll be happy to e-mail you the results of the study as soon as we complete it.